JUILLET 1987

HETEROGENEITY AND HAZARD DOMINANCE

IN DURATION DATA MODELS

n° 8736

Fourgeaud, C.(*), Gouriéroux, C.(*), and J. Pradel(**)

(*) CEPREMAP

(**) Paris IX University

HETEROGENEITY AND HAZARD DOMINANCE IN DURATION DATA MODELS

Abstract

We present some properties concerning heterogeneity using as illustrations duration data models. We first study the links between aggregate and disaggregate concepts. This allows to give some precise definitions of the heterogeneity biases and of the mover stayer phenomenon. We also introduce some natural orderings on the distributions of duration variables and on the heterogeneity distributions and we examine the effect on the aggregate duration variable of a "greater heterogeneity".

HETEROGENEITE ET DOMINANCE STOCHASTIQUE DANS LES MODELES DE DUREE

Résumé

Nous présentons diverses propriétés des modèles avec hétérogénéité prenant comme illustration les modèles de durée. Nous étudions d'abord les liens entre les concepts agrégés et désagrégés. Ceci nous permet de donner des définitions précises des notions de biais d'hétérogénéité et du phénomène "mobile-stable". Nous introduisons également des relations d'ordre naturelles sur les distributions de variables de durée et sur les distributions d'hétérogénéité et nous examinons l'effet sur la variable de durée agrégée d'une "plus grande" hétérogénéité.

Keywords : heterogeneity, mover-stayer phenomenon, stochastic dominance.

Mots clefs : hétérogénéité, modèles de durée, dominance stochastique.

Nomenclature JEL : 211.

1) INTRODUCTION

In the usual linear model : $Y = Xb + u$, the error term is often considered as an aggregate of various random variables with a number of different interpretations. One component of the disturbance summarizes the random feature of the observed phenomenon, but other components correspond to omitted variables (orthogonal to the introduced ones), to the non constancy of parameter b over time or individuals .. These last components are often called heterogeneity components.

If, in such a linear model, the truly random component and the heterogeneity one may be added to give the disturbance term, such a simple aggregation is generally difficult to justify in a non linear context. The omission of the heterogeneity component in a linear model, which is not in practice first order identifiable, has no consequence on simple estimation procedure such as least square method. However it may be very important to take into account this component in more complex models for either the consistency of the usual estimation procedures or the interpretation of the results.

In this paper we present some properties concerning heterogeneity using as illustrations duration data models. In the ten last years these models have been extensively considered, in relation with the increase of unemployment [see e.g. Nickell (1979), Kalbfleisch-Prentice (1980), Lancaster-Nickell (1980), Flinn-Heckman (1982), Lancaster-Chesher (1983), Heckman-Singer (1984) ...].

In section 2, we describe a duration data model with heterogeneity, both at disaggregate and aggregate levels (i.e. after integrating out the heterogeneity factor). This study allows us to easily relate the results at the aggregate level with the average of disaggregate results. In general these two notions do not coincide and this points out the existence of some heterogeneity bias. The sign of this bias is determined in different cases in particular concerning estimation of the hazard function and the expected residual life. This bias is often explained in the literature by the so-called Mover-stayer phenomenon (considering the simple case of a desaggregate exponential model). In this paper we give a precise description of this phenomenon by looking at the evolution over time of the heterogeneity distributions.

In section 3, we introduce some orderings on the distributions of duration variables and on the heterogeneity distributions. Then we examine the effect on the aggregate duration variable of a "greater heterogeneity".

Some technical proofs are gathered in appendices.

2) THE MODEL WITH HETEROGENEITY AND SOME OF ITS PROPERTIES

2.a Distribution of a duration variable

The main characteristic of a duration variable τ is to be a positive continuous random variable. Its distribution may of course be described by mean of its density function f or of its cumulative distribution function F. It may also be characterised using some other functions, which are easy to interprete and often introduce naturally in a number of computations. Among these other functions the most important ones are the survival function, the hazard function and the expected residual life [see Cox-Oakes (1984)]. The survival function is :

(1) $$S(t) = 1 - F(t) \ ;$$

it gives the probability for variable τ to be larger than the given value t : $S(t) = P(\tau > t)$. This is a continuous, non increasing function satisfying $S(0) = 1$ and $S(+\infty) = 0$.

The hazard function (or risk function) is defined by :

(2) $$\lambda(t) = \frac{f(t)}{1-F(t)} = \frac{f(t)}{S(t)} = -\frac{d}{dt} \operatorname{Log} S(t) \; .$$

If τ is interpreted as the duration spent in a given state, the value $\lambda(t)$ can be viewed as the instantaneous rate of leaving this state, since :

$$\lambda(t) = \lim_{dt \mapsto 0^+} \frac{1}{dt} P[\, t < \tau < t + dt \,/\, \tau > t \,] \, .$$

The hazard function characterizes the distribution of τ, since we have the relation :

(3) $$S(t) = \exp \left[- \int_0^t \lambda(u)\, du \right] .$$

Finally let us define the expected residual life. If at time t, an individual is alive, and if τ is the date of his death, his residual life is $\tau - t$. Taking the expectation, we obtain :

(4) $$r(t) = E\, [\, \tau - t \,/\, \tau > t \,] \, .$$

This function is also in one to one relationship with S and we have [Kalbfleisch-Prentice (1980) p 7] :

(5) $$r(t) = \frac{1}{S(t)} \int_t^\infty S(u)\, du \; ,$$

(6) $$S(t) = \frac{r(0)}{r(t)} \exp \left[- \int_0^t \frac{du}{r(u)} \right]$$

2.b Heterogeneity

From now on, we are interested in a population of individuals [e.g., unemployed people] with different characteristics described by mean of an index $\theta \in \Theta \subset \mathbb{R}$. The whole population $\mathcal{P}$ may be partitioned into the sub-populations $\mathcal{P}_\theta$, containing the individuals with index θ. The sizes of these sub-populations relative to the whole population are modelised by a p.d.f $\Pi(d\theta)$ defined on Θ. Finally we assume homogeneity within each sub-population : we consider that the (unemployment) spells of two individuals of a same sub-population $\mathcal{P}_\theta$ have the same distribution. At the disaggregate level of sub-population $\mathcal{P}_\theta$, the duration distribution is described by any one of the function :

$S(t\ ;\ \theta)$ [disaggregate survival function],

$\lambda(t\ ;\ \theta)$ [disaggregate hazard function],

$r(t\ ;\ \theta)$ [disaggregate expected residual life].

Throughout sub-populations, the value of θ is modified and also the duration distribution is ; parameter θ may be interpreted as an heterogeneity factor and Π as the heterogeneity distribution at initial date $t = 0$.

It is important to note that the heterogeneity distribution of unemployed people changes over time. When time increases, some individuals leave the unemployment state. Since the proportion of individuals leaving this state is not the same for the various sub-populations, the heterogeneity distribution likely depends on t. It is denoted by $\Pi_t(d\theta)$. To precise this remark we may consider the limit case of a population $\mathcal{P}$ of infinite size. The proportion of individuals of $\mathcal{P}_\theta$ staying unemployed at t is given by $S(t\ ;\ \theta)$. Thus we directly deduce :

$$\Pi_t(d\theta) = \frac{S(t\ ;\ \theta)\ \Pi(d\theta)}{\int_\Theta S(t\ ;\ \theta)\ \Pi(d\theta)}. \tag{6}$$

In the next subsections, E_t, V_t, Cov_t are the expectation, variance and covariance with respect to this distribution.

2.c Relation between disaggregate and aggregate statistics

Thus far we have described the duration at the desaggregate level. Now we can consider the distribution of this variable at the aggregate level, i.e. corresponding to an individual randomly sampled in population $\mathcal{P}$. The aggregate survival function is deduced from the disaggregate ones by the direct formula :

$$S(t) = \int_{\Theta} S(t\ ;\ \theta)\ \Pi(d\theta)\ . \tag{7}$$

The two other functions are obtained by applying formulas (2) and (5) :

$$\lambda(t) = -\frac{d \operatorname{Log} S(t)}{dt}\ ,\ r(t) = \frac{1}{S(t)} \int_t^{\infty} S(u)\ du\ .$$

Property 8 : If it is possible to commute integration and derivation, we have :

$$\lambda(t) = E_t\ \lambda(t\ ;\ \theta),$$

$$r(t) = E_t\ r(t\ ;\ \theta).$$

Proof : See appendix 1.

The instantaneous rate of leaving only concerns individuals unemployed at time t. Therefore it is natural to apply an agregration formula with respect to the distribution Π_t associated with these individuals. The same remark applies for the expected residual life. These simple relations are the basis of a joint study of these statistics at the aggregate and disaggregate levels.

2.d Two preliminary lemmas

For this comparison we shall use the two following results, whose proofs are given in appendix 2.

Lemma 9 : If $g(t\ ;\ \theta)$ is a function from $\mathbb{R}^{+} \times \Theta$ on $\mathbb{R}$, which is differentiable with respect to t, we have :

$$\frac{\partial}{\partial t} \underset{t}{E}\, g(t\ ;\ \theta) = \underset{t}{E}\left[\frac{\partial g(t\ ;\ \theta)}{\partial t}\right] - \underset{t}{\mathrm{Cov}}\,[g(t\ ;\ \theta),\ \lambda(t,\ \theta)]\ .$$

Lemma 10 : If this function $g(t\ ;\ \theta)$ is twice differentiable with respect to t, we have :

$$\begin{aligned}\frac{\partial^2}{\partial t^2} \underset{t}{E}\, g(t\ ;\ \theta) &= \underset{t}{E}\left[\frac{\partial^2 g(t\ ;\ \theta)}{\partial t^2}\right] - 2\,\underset{t}{\mathrm{Cov}}\left[\frac{\partial g}{\partial t}(t\ ;\ \theta),\ \lambda(t\ ;\ \theta)\right] \\ &- \underset{t}{\mathrm{Cov}}\left[g(t\ ;\ \theta),\ \frac{\partial \lambda}{\partial t}(t\ ;\ \theta)\right] + \underset{t}{\mathrm{Cov}}\,[g(t\ ;\ \theta)\ \lambda(t\ ;\ \theta),\ \lambda(t\ ;\ \theta)] \\ &- \underset{t}{\mathrm{Cov}}\,[g(t\ ;\ \theta),\ \lambda(t\ ;\ \theta)]\ \underset{t}{E}\,\lambda(t\ ;\ \theta) - \underset{t}{E}\,g(t\ ;\ \theta)\ \underset{t}{V}\,\lambda(t\ ;\ \theta).\end{aligned}$$

2.e Compared evolutions of aggregate and disaggregate hazard functions

Since $\lambda(t) = \underset{t}{E}\,\lambda(t\ ;\ \theta)$, we can apply lemmas (9) and (10) with : $g(t\ ;\ \theta) = \lambda(t\ ;\ \theta)$; we obtain :

Property 11 : $$\frac{\partial}{\partial t}\lambda(t) = \frac{\partial}{\partial t}\underset{t}{E}\,\lambda(t\ ;\ \theta) = \underset{t}{E}\left[\frac{\partial}{\partial t}\lambda(t\ ;\ \theta)\right] - \underset{t}{V}\,[\lambda(t\ ;\ \theta)].$$

<u>Property 12</u> :

$$\frac{\partial^2}{\partial t^2}\lambda(t) = \underset{t}{E}\left[\frac{\partial^2}{\partial t^2}\lambda(t\ ;\ \theta)\right] - 3\ \underset{t}{\mathrm{Cov}}\left[\frac{\partial\lambda(t\ ;\ \theta)}{\partial t},\ \lambda(t\ ;\ \theta)\right]$$

$$+ \underset{t}{\mathrm{Cov}}\ [\lambda^2(t\ ;\ \theta),\ \lambda(t\ ;\ \theta)] - 2\ \underset{t}{E}\ [\lambda(t\ ;\ \theta)]\ \underset{t}{V}\ [\lambda(t\ ;\ \theta)].$$

The first relation : $\frac{\partial}{\partial t}\lambda(t) = \underset{t}{E}\left[\frac{\partial}{\partial t}\lambda(t\ ;\ \theta)\right] - \underset{t}{V}\ [\lambda(t\ ;\ \theta)]$ has an interpretation, which is well known in the simple case : $\lambda(t\ ;\ \theta) = \theta$ [see also corollary 13]. To "estimate" the evolution of the hazard function associated with the average individual of population $\mathcal{P}$, i.e. to evaluate $\frac{\partial}{\partial t}\lambda(t)$, it is natural to use the average evolution of the individual hazard functions, i.e., $\underset{t}{E}\left[\frac{\partial}{\partial t}\lambda(t\ ;\ \theta)\right]$. Equation (11) says that this approach leads to an overestimation, since :

$$\underset{t}{E}\left[\frac{\partial}{\partial t}\lambda(t\ ;\ \theta)\right] \geqslant \frac{\partial\lambda}{\partial t}(t).$$

Moreover this difference between the quantity to evaluate and its natural approximation only depends on the variability of the instantaneous rate of leaving the state.

The existence of such an heterogeneity bias is often explained in duration models by the so-called mover-stayer phenomenon. Among the individuals, some have larger instantaneous rate of leaving and therefore leave more easily the state ; this implies a decrease for the average probability of leaving corresponding to individuals staying in the state. If this mover-stayer property explains the heterogeneity bias, these two notions are not equivalent. In subsection 2.g we shall precise the mover-stayer effect by mean of the time dependence of the heterogeneity distributions Π_t.

Generally the heterogeneity bias is discussed in the literature assuming time independence at disaggregate level [see e.g, Heckman-Singer (1984) p. 77-78]. In such a case the result has a very simple form :

<u>Corollary 13</u> : If $\lambda(t\ ;\ \theta) = \theta$, we have $\dfrac{\partial\lambda(t)}{\partial t} \leqslant 0$.

A microeconomic time independence $\dfrac{\partial\lambda(t\ ;\ \theta)}{\partial t} = 0$ implies negative duration dependence at macro-level.

Property (12) also shows an heterogeneity bias on second order derivative, since : $\dfrac{\partial^2\ \lambda(t)}{\partial t^2} \neq \underset{t}{E} \left[\dfrac{\partial^2}{\partial t^2}\ \lambda(t\ ;\ \theta)\right]$; however this bias has not a fixed sign.

This bias may be precised when the duration model is exponential at the disaggregate level.

<u>Corollary 14</u> : If $\lambda(t\ ;\ \theta) = \theta$, we have :

$$\frac{\partial^2\ \lambda(t)}{\partial t^2} = \underset{t}{\mathrm{Cov}} \left[(\theta - \underset{t}{E}\ \theta)^2\ ,\ \theta - \underset{t}{E}\ \theta\right] = E(\theta - \underset{t}{E})^3$$

<u>Proof</u> : We have : $\dfrac{d^2\ \lambda(t)}{dt^2} = \underset{t}{\mathrm{Cov}}\ (\theta^2\ ,\ \theta) - 2\ \underset{t}{E}\ \theta\ \underset{t}{V}\ \theta$

$$= \underset{t}{\mathrm{Cov}} \left[(\theta - \underset{t}{E}\ \theta)^2\ ,\ \theta - \underset{t}{E}\ \theta\right].$$

Q.E.D.

Clearly the sign of the heterogeneity bias depends on the asymmetric character of the heterogeneity distribution Π_t.

2.f Compared evolutions of the aggregate and disaggregate residual lifes

This study is easy if we assume that the individual hazard functions can be classified in increasing order. Such a classification is only available if two hazard functions associated with different values of the heterogeneity factor θ do not intersect. We give below the mathematical form of this assumption.

Assumption A.1 : The parameter θ belongs to an interval $\Theta \subset \mathbb{R}$ and the partial mapping : $\theta \mapsto \lambda(t\ ;\ \theta)$ is non decreasing for any value of t.

Under this assumption it is also possible to rank the individual expected residual lifes.

Property 15 : Under A1, the partial mappings : $\theta \mapsto r(t\ ;\ \theta)$ are non increasing, for any t.

Proof : We deduce from formulas (3) and (5) :

$$r(t\ ;\ \theta) = \frac{1}{S(t\ ;\ \theta)} \int_t^\infty S(u\ ;\ \theta)\, du$$

$$= \int_t^\infty \frac{\exp - \int_0^u \lambda(v\ ;\ \theta)\, dv}{\exp - \int_0^t \lambda(v\ ;\theta)\, dv}\, du$$

$$= \int_0^\infty \exp \left(- \int_t^{t+u} \lambda(v\ ;\ \theta)\, dv \right) du\ .$$

Since the exponential function and the operator $\int .\ du$ are increasing, the result directly follows.

Q.E.D.

The larger is the probability of finding a job, the smaller is the average residual duration of the unemployment spell. This interpretation is of course very natural.

Note that in the time independence case property 15 is immediate, since $\lambda(t\ ;\ \theta) = \theta$ and $r(t\ ;\ \theta) = \frac{1}{\theta}$.

Since $r(t) = \underset{t}{E}\, r(t\ ;\ \theta)$, we can apply lemmas (9) and (10) with $g(t\ ;\ \theta) = r(t\ ;\ \theta)$.

<u>Property 16</u> : $\frac{dr(t)}{dt} = \underset{t}{E}\left[\frac{\partial}{\partial t} r(t\ ;\ \theta)\right] - \underset{t}{\mathrm{Cov}}\,[\,r(t\ ;\ \theta),\ \lambda(t\ ;\ \theta)\,]$.

<u>Proof</u> : This is a direct consequence of lemma (9).

Q.E.D.

At this stage it is interesting to remark that using (5) we have :

$$\frac{d}{dt} r(t\ ;\ \theta) = -1 - \frac{\frac{\partial}{\partial t} S(t\ ;\ \theta)}{S(t\ ;\ \theta)^2} \int_t^{\infty} S(u\ ;\ \theta)\, du$$

$$= -1 + r(t\ ;\ \theta)\ \lambda(t\ ;\ \theta).$$

At the aggregate level, property 16 gives the same formula, since :

$$\frac{d}{dt} r(t) = \underset{t}{E}\,[-1 + r(t\ ;\ \theta)\ \lambda(t\ ;\ \theta)] - \underset{t}{\mathrm{Cov}}\,[\,r(t\ ;\ \theta),\ \lambda(t\ ;\ \theta)\,]$$

$$= -1 + \underset{t}{E}\, r(t\ ;\ \theta)\ \underset{t}{E}\, \lambda(t\ ;\ \theta) = -1 + r(t)\ \lambda(t).$$

Formula (16) can be interpreted in term of heterogeneity bias, since $\frac{d}{d\,t} r(t) \neq \underset{t}{E}\left[\frac{\partial}{\partial\, t} r(t,\ \theta)\right]$. Moreover the sign of this bias is fixed in an important case.

<u>Property 17</u> : Under assumption A 1, the heterogeneity bias : $\frac{d}{d\,t} r(t) - \underset{t}{E}\, \frac{\partial}{\partial\, t} r(t\ ;\ \theta)$ is positive.

<u>Proof</u> : We have : $\frac{d}{d\,t} r(t) - \underset{t}{E}\, \frac{\partial}{\partial\, t} r(t\ ;\ \theta) = -\underset{t}{\mathrm{cov}}\,[r(t\ ;\ \theta),\ \lambda(t\ ;\ \theta)\,]$.

Under assumption A 1, we know that functions $- r(t\,;\,\theta)$ and $\lambda(t\,;\,\theta)$ are non decreasing in θ and therefore they are positively correlated [see appendix 4] :

$$\underset{\theta}{\mathrm{Cov}}\,(- r(t\,;\,\theta)\,,\,\lambda(t\,;\,\theta)) \geqslant 0.$$

The result follows.

Q.E.D.

This is a result similar to the one obtained for hazard functions : if the natural approximation of $\frac{\partial\lambda(t)}{dt}$ overestimate this quantity, the natural approximation of $\frac{d}{d\,t}\,r(t)$ leads to underestimation. The interpretation as a consequence of mover-stayer phenomenon is of course the same.

As usual the result takes a simpler form with time independence.

<u>Corollary 18</u> : If $\lambda(t\,;\,\theta) = \theta$ [$\Leftrightarrow r(t\,;\theta) = \frac{1}{\theta}$], the aggregate expected residual life is non decreasing : $\frac{dr(t)}{dt} \geqslant 0$.

It is also possible to examine second order derivative of function r. We give the result under time independence.

<u>Corollary 19</u> : If $\lambda(t\,;\,\theta) = \theta$ [$\Leftrightarrow r(t\,;\,\theta) = \frac{1}{\theta}$], we have :

$$\frac{d^2}{dt^2}\,r(t) = - \underset{t}{V}\,\theta\;\;\underset{t}{E}\,\theta\left\{\frac{\underset{t}{\mathrm{Cov}}\,(\frac{1}{\theta},\,\theta)}{\underset{t}{V}\,\theta} + \frac{\underset{t}{E}\,(\frac{1}{\theta})}{\underset{t}{E}\,\theta}\right\}.$$

Proof : We apply lemma (10) noting that the derivatives of $g(t\,;\,\theta) = r(t\,;\,\theta)$ and $\lambda(t\,;\,\theta)$ with respect to θ are equal to zero and using the equality : $g(t\,;\,\theta)\ \lambda(t\,;\,\theta) = 1$. We obtain :

$$\frac{d^2 r(t)}{dt^2} = - \operatorname{Cov}_t \left(\frac{1}{\theta}, \theta\right) E_t\, \theta - E_t \left(\frac{1}{\theta}\right) V_t\, \theta$$

$$= - V_t(\theta)\ E_t(\theta) \left[\frac{\operatorname{Cov}_t (\frac{1}{\theta}, \theta)}{V_t\, \theta} + \frac{E_t (\frac{1}{\theta})}{E_t\, \theta}\right].$$

Q.E.D.

For discussing the sign of the second derivative, we have to compare the regression coefficient of $\frac{1}{\theta}$ on θ with the slope of the line joining the origine to the point $(\theta, \frac{1}{\theta})$. It is easily seen that even under time independence the sign is not fixed.

2.g Time evolution of the heterogeneity distribution

Finally let us now compare the heterogeneity distributions at two different dates.

Property 20 : Under assumption A1, we have : $E_t\, g(\theta) \geq E_{t'}\, g(\theta)$, for any non decreasing function g and for any pair of real numbers $t' \geq t$.

Proof : Let us apply lemma (9) with $g(t\,;\,\theta) = g(\theta)$. We obtain :

$$\frac{\partial}{\partial t} E_t\, g(\theta) = E_t \left[\frac{\partial}{\partial t} g(\theta)\right] - \operatorname{Cov}_t\, [g(\theta), \lambda(t\,;\,\theta)]$$

$$= - \operatorname{Cov}_t\, [\, g(\theta)\ ,\ \lambda(t\,;\,\theta)\,]\,.$$

This last quantity is non positive, since g and $\lambda(t\,;\,.)$ are non decreasing functions of θ [see appendix 4]. We deduce that the mapping : $t \to E_t\, g(\theta)$ is non increasing.

Q.E.D.

The condition of property 20 is the usual stochastic dominance condition at order one [see, e.g. Brumelle-Vickson (1975)]. Therefore the condition can also be expressed using the c.d.f. :

$$Q_t(\theta) = \int_0^\theta \Pi_t(dv) .$$

<u>Corollary 21</u> : Under assumption A1 and if $t' > t$, Π_t dominates $\Pi_{t'}$ at order one or equivalently : $\forall\ \theta,\ Q_{t'}(\theta) \geq Q_t(\theta)$.

This last inequality is clearly the mathematical description of the mover-stayer phenomenon. With as time increases, the proportion of stayers increases and this result is valid for any possible definition of what is a stayer, i.e., for any value of θ giving the boundary between mover and stayer.

The sequence of cumulative distribution functions Q_t is non decreasing, and, since θ is non negative, we directly deduce weak convergence of Π_t to a limit distribution Π_∞.

<u>Corollary 22</u> : Under assumption A 1, the heterogeneity distribution Π_t weakly converges to a limit distribution Π_∞, when t tends to infinity.

The limit Π_∞ may be considered as a minimum degree of heterogeneity. In some cases this minimum corresponds to homogeneity.

<u>Property 23</u> : Under time independence the limit distribution is a print mass.

<u>Proof</u> : i) If $\lambda(t\ ;\ \theta) = \theta$, the survival function is $S(t\ ;\ \theta) = \exp - \theta t$ and the heterogeneity distribution is

given by : $\Pi_t\,(d\theta) = \dfrac{\exp - \theta t\ \Pi(d\theta)}{\int_{\Theta} \exp - \theta t\ \Pi(d\theta)}$.

We deduce after time translation :

$$\Pi_{t+t_o}\,(d\theta) = \frac{\exp - \theta(t + t_o)\ \Pi(d\theta)}{\int_{\Theta} \exp - \theta(t + t_o)\ \Pi(d\theta)}$$

$$= \frac{\exp - \theta t \dfrac{\exp - \theta t_o\ \Pi(d\theta)}{\int_{\Theta} \exp - \theta t_o\ \Pi(d\theta)}}{\int_{\Theta} \exp - \theta\, t \dfrac{\exp - \theta t_o}{\int_{\Theta} \exp - \theta t_o\ \Pi(d\theta)}\ \Pi(d\theta)}$$

$$= \frac{\exp - \theta t\ \Pi_{t_o}\,(d\theta)}{\int_{\Theta} \exp - \theta t\ \Pi_{t_o}(d\theta)} .$$

ii) Therefore for any bounded continuous function g :

$$\underset{t+t_o}{E}\ g(\theta) = \underset{t_o}{E}\ [\exp - \theta t\ g(\theta)] \Big/ \underset{t_o}{E}\ (\exp - \theta t) ,$$

and, when t_o tends to infinity, this relation becomes :

$$\underset{\infty}{E}\ g(\theta)\ \underset{\infty}{E}\ (\exp - \theta t) = \underset{\infty}{E}\ [\exp - \theta t\ g(\theta)] \ , \ \forall\, t, \ \ \forall\, g.$$

iii) Choosing $g(\theta) = \exp - \theta\tau$, we have :

$$\underset{\infty}{E}\ (\exp - \theta\tau)\ \underset{\infty}{E}\ (\exp - \theta t) = \underset{\infty}{E}\ \exp - \theta(t + \tau) \ , \ \forall\, t, \ \tau.$$

Thus the Laplace transform $\mathcal{L}_\infty(t) = \underset{\infty}{E}(\exp - \theta t)$ of the limit distribution Π_∞ satisfies :

$$\mathcal{L}_\infty(\tau)\ \mathcal{L}_\infty(t) = \mathcal{L}_\infty(t+\tau) \quad , \forall\, t, \tau$$

This implies the specific form $\mathcal{L}_\infty(t) = e^{at}$ for some positive real number a and the equality of Π_∞ with the ponctual mass at a.

Q.E.D.

A second order effect on the heterogeneity distribution Π_t may also be noted.

<u>Property 24</u> Under assumption A 1, if the initial distribution is absolutely continuous $\Pi(d\theta) = q(d\theta).\ d\theta$ with a differentiable p.d.f and if the hazard function $\lambda(t\ ;\ \theta)$ is jointly differentiable with respect to t, θ, then the heterogeneity distribution $\Pi_t\ (d\theta)$ are continuous with p.d.f. q_t satisfying :

$$\frac{\partial \operatorname{Log} q_{t'}(\theta)}{\partial \theta} \leqslant \frac{\partial \operatorname{Log} q_t(\theta)}{\partial \theta} \quad , \quad \forall\, \theta\ , \ \forall\, t' > t.$$

<u>Proof</u> : We have :

$$q_t(\theta) = \frac{S(t\ ;\ \theta)\ q(\theta)}{\int_\Theta S(t\ ;\ \theta)\ q(\theta)\ d\theta},$$

and : $\operatorname{Log} q_t(\theta) = \operatorname{Log} S(t\ ;\ \theta) + \operatorname{Log} q(\theta) - \operatorname{Log} \int_\Theta S(t\ ;\ \theta)\ q(\theta)\ d\theta.$

By differentiation with respect to θ,t, we deduce :

$$\begin{aligned} \frac{\partial^2 \operatorname{Log} q_t(\theta)}{\partial t\, \partial \theta} &= \frac{\partial^2 \operatorname{Log} S(t\ ;\ \theta)}{\partial t\, \partial \theta} \\ &= - \frac{\partial}{\partial \theta} \frac{\partial}{\partial t} [- \operatorname{Log} S(t\ ;\ \theta)] \\ &= - \frac{\partial}{\partial \theta} \lambda(t\ ;\ \theta). \end{aligned}$$

Since, under assumption A 1, the hazard function $\lambda(t\ ;\ \theta)$ is non decreasing with respect to θ, we have :

$$\frac{\partial}{\partial \theta} \lambda(t ; \theta) \geq 0 \Leftrightarrow \frac{\partial^2 \text{Log } q_t(\theta)}{\partial t \, \partial \theta} \leq 0.$$

Q.E.D.

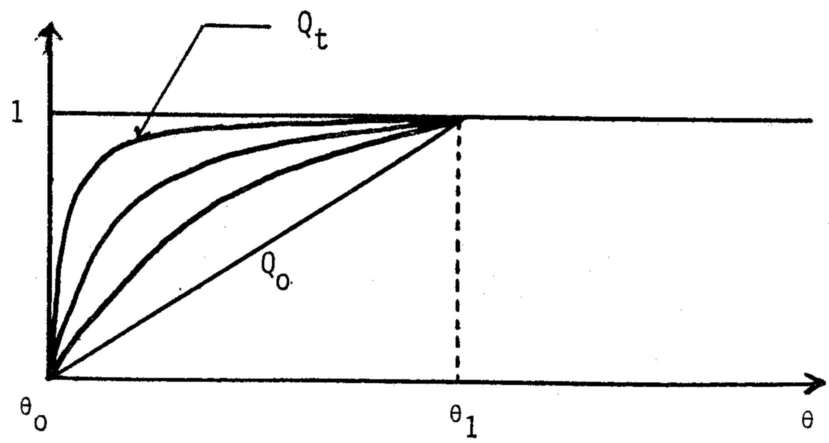
<u>Corollary 25</u> : Under the assumptions of property (24) and if the initial heterogeneity distribution is the uniform distribution on $[\theta_o, \theta_1]$, then the c.d.f. Q_t is a concave function.

<u>Proof</u> : Since : $\dfrac{\partial \text{ Log } q_t(\theta)}{\partial \theta} \leq \dfrac{\partial \text{ Log } q_o(\theta)}{\partial \theta} = 0$, we directly deduce :

$$\frac{\partial q_t(\theta)}{\partial \theta} = \frac{\partial^2 Q_t(\theta)}{\partial \theta^2} \leq 0.$$

Q.E.D.

The sequence of heterogeneity c.d.f. has the following representation under the assumptions of the previous corollary.



Another illustration of the previous results is obtained by mixing a time independent model : $S(t ; \theta) = \exp - \theta t$, with a gamma heterogeneity distribution :

$$q(\theta) = \frac{a^{\nu} \theta^{\nu-1} \exp - a\theta}{\Gamma(\nu)}.$$

The heterogeneity p.d.f. at time t is :

$$q_t(\theta) = \frac{\exp - \theta t \, \dfrac{a^{\nu}\,\theta^{\nu-1} \exp - a\theta}{\Gamma(\nu)}}{\displaystyle\int_0^{\infty} \exp - \theta t \, \frac{a^{\nu}\,\theta^{\nu-1} \exp - a\theta}{\Gamma(\nu)}\, d\theta}$$

$$= \frac{(a+t)^{\nu-1} \exp - \theta(a+t)\ \theta^{\nu-1}}{\Gamma(\nu)} .$$

At date t the distribution of $\dfrac{\theta}{a+t}$ is a gamma distribution with parameter ν. The c.d.f. are thus modified and, at the limit $t \to \infty$, we have the convergence to a ponctual mass at zero.

Since $\dfrac{\partial}{\partial \theta} \text{Log } q_t(\theta) = -(a+t) + \dfrac{\nu - 1}{\theta}$ is non increasing in t, we immediately verify property 24. However it is known that gamma distributions have a non monotonic density function and corollary 25 cannot be extended to this case.

3) <u>ORDERING ON DURATION DISTRIBUTIONS</u>

To precise the possible interpretations of parameters appearing in duration data models, it is interesting to introduce different orderings to classify the distributions of positive random variables. For illustration we consider that τ is the lifetime of an individual and we choose some binary relations to describe the preference for living during a long time.

The same binary relations may also be used for study on unemployment ; they have to be transformed by symmetry, since the individual has a preference for short unemployment spells. There exist several natural definitions of such an ordering. We give them successively, we describe the links between these definitions and we relate these orderings to the notion of heterogeneity.

3.a <u>Stochastic dominance of order 1</u>

The most usual ordering on probability distributions is defined in the following way.

<u>Definition 26</u> : Let us consider two possible duration distributions P^* and P with respective p.d.f. f^* and f. P^* dominates P at order one if and only if for any positive non decreasing function V, we have :

$$\overset{*}{E}\, V(\tau) = \int_0^\infty V(t)\, f^*(t)\, dt \geqslant E\, V(\tau) = \int_0^\infty V(t)\, f(t)\, dt.$$

Function V may be interpreted as an utility function and the previous condition describes a preference for large values of duration τ, which is uniform in the utility functions V. As previously mentioned this ordering can be expressed using the survival functions :

<u>Property 27</u> : P^* dominates P at order 1 iff :

$$S^*(t) \geqslant S(t)\ , \quad \forall\, t.$$

The condition is satisfied iff the probability to attain a given age t is larger with P^* and this for any value of this age.

Now we are going to consider duration models with heterogeneity, to introduce the idea of a large heterogeneity and to examine the implications on the survival functions. To simplify the presentation, we restrict ourselves to the case of time independence.

The first level of heterogeneity is obtained by introducing an heterogeneity factor $\lambda(t\,;\theta) = \theta$, whose distribution is $\Pi_1(d\theta)$.

The second level of heterogeneity is deduced from the first one by the addition of another heterogeneity factor α :

$$\lambda(t\ ;\ \theta,\ \alpha) = \theta + \alpha\ .$$

These factors are such that the marginal distribution of θ is $\Pi_1(d\theta)$ and the conditional distribution of α given θ, $\Pi_2(d\alpha/\theta)$ is zero mean. Clearly the second model is more heterogenous than the first one.

The two models may be described equivalently in the following way :

(*) at the first level, we have : $\lambda(t\ ;\ \theta) = \theta$ with a distribution $\Pi_1(d\theta)$,

(**) at the second level, we have : $\lambda(t\ ;\ \theta) = \theta$, with a distribution $\Pi_1^*(d\theta)$ defined by :

$$\int_\Theta g(\theta)\ \Pi_1^*(d\theta) = \iint_{\Theta^2} g(\theta + \alpha)\ \Pi_2(d\alpha\ /\ \theta)\ \Pi_1(d\theta).$$

With this formulation we see that the idea of a larger heterogeneity is directly related to the usual second order stochastic dominance [see Rothschild - Stiglitz (1970), (1973), Brumelle - Vickson (1975)].

<u>Definition 28</u> : The distribution Π_1^* gives a larger heterogeneity than Π_1 does iff Π_1^* dominates Π_1 at second order.

Now we can determine the aggregate survival functions associated with the heterogeneity distribution Π_1 and Π_1^*.

We have : $S(t) = \int_\Theta \exp\ (-\ \theta\ t)\ \Pi_1(d\ \theta)$,

$$S^*(t) = \iint_{\Theta^2} \exp\ [-\ (\theta + \alpha)\ t]\ \Pi_2(d\alpha\ /\ \theta)\ \Pi_1(d\ \theta).$$

By using conditional expectations, we obtain :

$$S^*(t) = E\,E[\exp[-(\theta+\alpha)\,t] / \theta]$$

$$= E\,\{\exp(-\theta\,t)\,E(\exp - \alpha\,t / \theta)\}$$

$$\geq E\,\{\exp - \theta\,t \quad \exp[-E(\alpha / \theta)\,t]\}$$

(from the convexity inequality)

$$= E\,[\exp - \theta\,t] = S(t)$$

(since $E(\alpha / \theta) = 0$).

<u>Property 29</u> : Under time independence assumption, if Π_1^* gives a larger heterogeneity than Π_1, then P^* dominates P at order one.

3.b <u>Hazard dominance</u>

Another natural definition of an ordering on duration distributions is based on the hazard function. Recall that if τ is the lifetime, $\lambda(t)$ is the mortality rate at age t.

<u>Definition 30</u> : P^* hazard dominates P iff $\lambda(t) \geq \lambda^*(t), \quad \forall\, t$.

The link between the two orderings we have introduced is given below.

<u>Property 31</u> : P^* hazard dominates P iff for any positive non decreasing function V and for any positive real number t_o, we have : $\overset{*}{E}\,[V(\tau) / \tau \geq t_o] \geq E(V(\tau) / \tau \geq t_o]$.

<u>Proof</u> : see appendix 3.

The stochastic dominance of order 1 corresponds to the particular case, in which the relation is only considered at birth date $t = 0$.

Corollary 32 | If P^* hazard dominates P, then P^* dominates P at order 1.

The hazard dominance condition is more severe than the stochastic dominance at order 1. Therefore it is also necessary to introduce an adapted ordering on the heterogeneity distributions. With the same notations as in the previous subsection, the hazard functions are given by :

$$\lambda(t) = \frac{\int_\Theta \theta \exp - \theta t \; \Pi_1(d\theta)}{\int_\Theta \exp - \theta t \; \Pi_1(d\theta)} = \frac{E[\theta \exp - \theta t]}{E[\exp - \theta t]},$$

$$\lambda^*(t) = \frac{\iint_{\Theta^2} (\theta + \alpha) \exp [- (\theta + \alpha) t] \, \Pi_2(d\alpha/\theta) \; \Pi_1(d\theta)}{\iint_{\Theta^2} \exp [- (\theta + \alpha) t] \, \Pi_2(d\alpha/\theta) \; \Pi_1(d\theta)}$$

$$= \frac{E [(\theta + \alpha) \exp - \theta t \exp - \alpha t]}{E [\exp - \theta t \exp - \alpha t]}$$

$$= \frac{E [\theta \exp - \theta t \; E (\exp - \alpha t/\theta)]}{E [\exp - \theta t \; E (\exp - \alpha t/\theta)]} + \frac{E \{\exp - \theta t \; E [\alpha \exp - \alpha t/\theta]\}}{E [\exp - \theta t \; E (\exp - \alpha t/\theta)]}.$$

Definition 33 : Π_1^* is strongly more heterogenous than Π_1 iff Π_1^* dominates Π_1 at second order and if $E[\exp - \alpha t/\theta]$ is a non increasing function in θ.

Property 34 : Under time independence, if Π_1^* is strongly more heterogenous than Π_1, then P^* hazard dominates P.

Proof : i) We have : $E[\alpha \exp - \alpha t/\theta] - E(\alpha/\theta) \; E[\exp - \alpha t/\theta] \leq 0$, since α and $\exp - \alpha t$ are respectively increasing and decreasing functions of α. Moreover since $E(\alpha/\theta) = 0$, we have $E[\alpha \exp - \alpha t/\theta] \leq 0$.

We deduce that :

$$\lambda^*(t) \leq \frac{E\,\{\theta \exp - \theta t\, E[\exp - \alpha t/\theta]\}}{E\,\{\exp - \theta t\, E(\exp - \alpha t/\theta]\}}\,.$$

ii) On the other hand, since $E[\exp - \alpha t/\theta]$ is non increasing with respect to θ, we have :

$$\underset{t}{\mathrm{Cov}}\,(\theta,\, E(\exp - \alpha t/\theta)) \leq 0$$

$$\longleftrightarrow \frac{E\,[\,\theta \exp - \theta t\, E(\exp - \alpha t/\theta)\,]}{E(\exp - \theta t)} - \frac{E(\theta \exp - \theta t)}{E(\exp - \theta t)}\,\frac{E\,[\exp - \theta t\, E(\exp - \alpha t/\theta)\,]}{E(\exp - \theta t)} \leq 0\,,$$

$$\longleftrightarrow \frac{E\,[\,\theta \exp - \theta t\, E(\exp - \alpha t/\theta)\,]}{E\,[\exp - \theta t\, E(\exp - \alpha t/\theta)\,]} \leq \frac{E(\theta \exp - \theta t)}{E(\exp - \theta t)}\,.$$

This implies : $\lambda^*(t) \leq \lambda(t)$.

Q.E.D.

4) CONCLUDING REMARKS

In the previous sections, we have described some probabilistic properties of models with heterogeneity. Such a study has to be considered as a first step before considering the consequences of heterogeneity for statistical inference. For instance the following statistical problems have not been completely solved in the previous leterature.

i) Let us consider a model with heterogeneity, with a disaggregate hazard function $\lambda(t \; ; b \; ; \theta)$ depending on a parameter b and with a given heterogeneity distribution Π_o. The asymptotic covariance matrix of the maximum likelihood estimator of b depends on Π_o. Does a larger heterogeneity implies a worse precision on b ?

ii) If the previous model is estimated under the assumption of homogeneity, i.e. choosing $\theta = \theta_o$, the constrained M.L. estimator of b is in general unconsistent if heterogeneity is present. What is the precise dependence between the heterogeneity bias and the distribution b.

iii) Does the Lagrange multiplier statistic of the homogeneity hypothesis give a natural measure of the degree of heterogeneity ?

APPENDIX 1

Relations between aggregate and desaggregate statistics

Property | $\lambda(t) = \underset{t}{E}\ \lambda(t\ ;\ \theta)$.

We have :

$$\lambda(t) = -\frac{\partial \operatorname{Log} S(t)}{\partial t} = -\frac{\partial}{\partial t} \operatorname{Log} \int_{\Theta} S(t\ ;\ \theta)\ \Pi(d\theta)$$

$$= -\frac{\dfrac{\partial}{\partial t}\displaystyle\int_{\Theta} S(t\ ;\ \theta)\ \Pi(d\theta)}{\displaystyle\int_{\Theta} S(t\ ;\ \theta)\ \Pi(d\theta)}$$

$$= \frac{\displaystyle\int_{\Theta} -\frac{\partial S(t\ ;\ \theta)}{\partial t}\ \Pi(d\theta)}{\displaystyle\int_{\Theta} S(t\ ;\ \theta)\ \Pi(d\theta)}$$

$$= \int_{\Theta} -\frac{\partial \operatorname{Log} S(t\ ;\ \theta)}{\partial t}\ \frac{S(t\ ;\ \theta)\ \Pi(d\theta)}{\displaystyle\int_{\Theta} S(t\ ;\ \theta)\ \Pi(d\theta)}$$

$$= \int_{\Theta} \lambda(t\ ;\ \theta)\ \Pi_t(d\theta) = \underset{t}{E}\ \lambda(t\ ;\ \theta)$$

Property | $r(t) = \underset{t}{E}\ r(t\ ;\ \theta)$.

Similary, we have :

$$r(t) = \frac{1}{S(t)} \int_t^{\infty} S(u)\ du$$

$$= \frac{1}{\displaystyle\int_{\Theta} S(t\ ;\ \theta)\ \Pi(d\theta)} \int_t^{\infty} \left[\int_{\Theta} S(u\ ;\ \theta)\ \Pi(d\theta) \right] du$$

$$= \frac{1}{\int_{\Theta} S(t\ ;\ \theta)\ \Pi(d\theta)} \int_{\Theta} \left[\int_t^{\infty} S(u\ ;\ \theta)\ du \right] \Pi(d\theta)$$

$$= \frac{1}{\int_{\Theta} S(t\ ;\ \theta)\ \Pi(d\theta)} \int_{\Theta} r(t\ ;\ \theta)\ S(t\ ;\ \theta)\ \Pi(d\theta)$$

$$= \int_{\Theta} r(t\ ;\ \theta)\ \Pi_t(d\theta) = \underset{t}{E}\ r(t\ ;\ \theta).$$

APPENDIX 2

Proof of the preliminary lemmas

First derivative : $\frac{\partial}{\partial t} \underset{t}{E}\, g(t\ ;\ \theta)$.

We have :

$$\frac{\partial}{\partial t} \underset{t}{E}\, g(t\ ;\ \theta) = \frac{\partial}{\partial t}\left[\frac{\int_{\Theta} g(t\ ;\ \theta)\ S(t\ ;\ \theta)\ \Pi(d\theta)}{\int_{\Theta} S(t\ ;\ \theta)\ \Pi(d\theta)}\right]$$

$$= \frac{\int_{\Theta} \frac{\partial}{\partial t} g(t\ ;\ \theta)\ S(t\ ;\ \theta)\ \Pi(d\theta)}{\int_{\Theta} S(t\ ;\ \theta)\ \Pi(d\theta)} + \frac{\int_{\Theta} g(t\ ;\ \theta)\ \frac{\partial}{\partial t} S(t\ ;\ \theta)\ \Pi(d\theta)}{\int_{\Theta} S(t\ ;\ \theta)\ \Pi(d\theta)}$$

$$- \frac{\int_{\Theta} g(t\ ;\ \theta)\ S(t\ ;\ \theta)\ \Pi(d\theta)\ \int_{\Theta} \frac{\partial}{\partial t} S(t\ ;\ \theta)\ \Pi(d\theta)}{\left[\int_{\Theta} S(t\ ;\ \theta)\ \Pi(d\theta)\right]^2}$$

Therefore :

$$\frac{\partial}{\partial t} \underset{t}{E}\, g(t\ ;\ \theta) = \underset{t}{E}\left[\frac{\partial}{\partial t} g(t\ ;\ \theta)\right] + \underset{t}{E}\left[g(t\ ;\ \theta)\ \frac{\partial \operatorname{Log} S(t\ ;\ \theta)}{\partial t}\right]$$

$$- \underset{t}{E}\,[\,g(t\ ;\ \theta)\,]\ \underset{t}{E}\left[\frac{\partial \operatorname{Log} S(t\ ;\ \theta)}{\partial t}\right]$$

$$= \underset{t}{E}\left[\frac{\partial}{\partial t} g(t\ ;\ \theta)\right] - \underset{t}{E}\,[\,g(t\ ;\ \theta)\ \lambda(t\ ;\ \theta)\,] + \underset{t}{E}\,[\,g(t\ ;\ \theta)\,]\ \underset{t}{E}\,[\,\lambda(t\ ;\ \theta)\,]$$

$$= \underset{t}{E}\left[\frac{\partial}{\partial t} g(t\ ;\ \theta)\right] - \underset{t}{\operatorname{Cov}}\,[\,g(t\ ;\ \theta)\ ,\ \lambda(t\ ;\ \theta)\,]$$

Determination of the second derivative

$$\frac{\partial^2}{\partial t^2} \underset{t}{E}\, g(t\ ;\ \theta) = \frac{\partial}{\partial t} \underset{t}{E} \left[\frac{\partial}{\partial t} g(t\ ;\ \theta) \right] - \frac{\partial}{\partial t} \underset{t}{E}\, [\, g(t\ ;\ \theta)\ \lambda(t\ ;\ \theta)\,]$$

$$+ \frac{\partial}{d\, t} \left(\underset{t}{E}\, [\, g(t\ ;\ \theta)\,] \right) \underset{t}{E}\, \lambda(t\ ;\ \theta) + \underset{t}{E}\, g(t\ ;\ \theta)\ \frac{\partial}{d\, t} \left[\underset{t}{E}\, \lambda(t\ ;\ \theta) \right]$$

$$= \left[\underset{t}{E} \left[\frac{\partial^2}{\partial t^2} g(t\ ;\ \theta) \right) \right] - \underset{t}{\mathrm{Cov}} \left[\frac{\partial}{\partial t} g(t\ ;\ \theta),\ \lambda(t\ ;\ \theta) \right]$$

$$- \left\{ \underset{t}{E} \left[\frac{\partial}{\partial t} g(t\ ;\ \theta)\ \lambda(t\ ;\ \theta) \right] + \underset{t}{E} \left[g(t\ ;\ \theta)\ \frac{\partial}{d\, t} \lambda(t\ ;\ \theta) \right] \right.$$

$$\left. - \underset{t}{\mathrm{Cov}}\ [\, g(t\ ;\ \theta)\ \lambda(t\ ;\ \theta),\ \lambda(t\ ;\ \theta)\,] \right\}$$

$$+ \left\{ \underset{t}{E} \left[\frac{\partial}{d\, t} g(t\ ;\ \theta) \right] - \underset{t}{\mathrm{Cov}}\ [\, g(t\ ;\ \theta),\ \lambda(t\ ;\ \theta)\,] \right\} \underset{t}{E}\, \lambda(t\ ;\ \theta)$$

$$+ \underset{t}{E}\, g(t\ ;\theta) \left\{ \underset{t}{E} \left[\frac{\partial}{\partial t} \lambda(t\ ;\ \theta) \right] - \underset{t}{\mathrm{Cov}}\ [\, \lambda(t\ ;\ \theta),\ \lambda(t\ ;\ \theta)\,] \right\}$$

$$= \underset{t}{E} \left[\frac{\partial^2}{\partial t^2} g(t\ ;\ \theta) \right] - \underset{t}{\mathrm{Cov}} \left[\frac{\partial}{\partial t} g(t\ ;\ \theta),\ \lambda(t\ ;\ \theta) \right]$$

$$- \underset{t}{E} \left\{ \frac{\partial}{d\, t} [\, g(t\ ;\ \theta)\,]\ \lambda(t\ ;\ \theta) \right\}$$

$$- \underset{t}{E}\left[g(t\,;\,\theta)\,\frac{\partial}{d\,t}\,\lambda(t\,;\,\theta)\right] + \underset{t}{\mathrm{Cov}}\,[\,g(t\,;\,\theta)\;\lambda(t\,;\,\theta),\;\lambda(t\,;\,\theta)\,]$$

$$+ \underset{t}{E}\left[\frac{\partial}{\partial\,t}\,g(t\,;\,\theta)\right] \underset{t}{E}\,\lambda(t\,;\,\theta) - \underset{t}{\mathrm{Cov}}\,[\,g(t\,;\,\theta),\;\lambda(t\,;\,\theta)\,]\,\underset{t}{E}\,\lambda(t\,;\,\theta)$$

$$+ \underset{t}{E}\,[\,g(t\,;\,\theta)\,]\,\underset{t}{E}\left[\frac{\partial}{\partial\,t}\,\lambda(t\,;\,\theta)\right] - \underset{t}{E}\,g(t\,;\,\theta)\;\underset{t}{V}\,[\,\lambda(t\,;\,\theta)\,]$$

$$= \underset{t}{E}\left[\frac{\partial^2}{\partial t^2}\,g(t;\theta)\right] - 2\,\underset{t}{\mathrm{Cov}}\left[\frac{\partial}{dt}\,g(t;\theta),\;\lambda(t;\theta)\right] - \underset{t}{\mathrm{Cov}}\left[g(t\,;\,\theta),\;\frac{\partial}{\partial t}\,\lambda(t;\theta)\right]$$

$$+ \underset{t}{\mathrm{Cov}}\,[\,g(t\,;\,\theta)\;\lambda(t\,;\,\theta),\;\lambda(t\,;\,\theta)\,] - \underset{t}{\mathrm{Cov}}\,[\,g(t\,;\,\theta),\;\lambda(t\,;\,\theta)\,]\,\underset{t}{E}\,[\,\lambda(t\,;\,\theta)\,]$$

$$- \underset{t}{E}\,[\,g(t\,;\,\theta)\,]\,\underset{t}{V}\,[\,\lambda(t\,;\,\theta)\,]\,.$$

APPENDIX 3

Proof of property 31

i) The condition $\overset{*}{E}(V(\tau) / \tau \geq t_o) \geq E(V(\tau) / \tau \geq t_o)$ for any non decreasing function is a first order stochastic dominance condition applied to the conditional distribution. It is equivalent to :

$$S^*(t / t_o) \geq S(t / t_o), \quad \forall t \geq 0$$

$$\text{with : } S(t / t_o) = P\left[\tau > t + t_o / \tau > t_o\right] = \frac{S(t + t_o)}{S(t)}$$

$$= \exp - \int_{t_o}^{t+t_o} \lambda(u)\, du,$$

and a similar expression for $S^*(t / t_o)$.

ii) Necessary part :

If $\lambda(t) \geq \lambda^*(t)$, $\forall t$, we also have : $\int_{t_o}^{t+t_o} \lambda(u)\, du \geq \int_{t_o}^{t+t_o} \lambda^*(u)\, du$,

and therefore : $S(t / t_o) \leq S^*(t / t_o)$, $\forall t \geq 0$, $\forall t_o \geq 0$.

iii) Sufficient part

Since : $S^*(t / t_o) \geq S(t / t_o)$, $\forall t \geq 0$, $\forall t_o \geq 0$, we also have :

$$\lim_{t \to 0} \frac{1}{t}\left[S^*(t / t_o) - 1\right] \geq \lim_{t \to 0} \frac{1}{t}\left[S(t / t_o) - 1\right]$$

The result is then deduced from the equality :

$$\lim_{t \to 0} \frac{1}{t}\left[S(t / t_o) - 1\right] = \lim_{t \to 0} \frac{1}{t}\left[\exp - \int_{t_o}^{t_o+t} \lambda(u)\, du - 1\right] = -\lambda(t_o).$$

Q.E.D.

APPENDIX 4

Correlation between two non decreasing functions of a same random variable

> If X is a random variable, g and h are two non decreasing functions, then $\mathrm{Cov}\ [\, g(x),\ h(x)\,] \geqslant 0$.

Let us consider two independent drawings X_1, X_2 from the distribution of X. Since g and h are non decreasing, we have :

$$[\, g(X_1) - g(X_2)\,]\ [\, h(X_1) - h(X_2)\,] \geqslant 0.$$

Taking the expectation we obtain :

$$E\ [\, g(X_1) - g(X_2)\,]\ [\, h(X_1) - h(X_2)\,] \geqslant 0$$

$\longleftrightarrow E\ [\, g(X_1)\ h(X_1)\,] - E\ g(X_2)\ E\ h(X_1) - E\ g(X_1)\ E\ h(X_2)$

$+ E\ [\, g(X_2)\ h(X_2)\,] \geqslant 0$ (using the independence between X_1 and X_2)

$\longleftrightarrow 2\ [\, E[\, g(X)\ h(X)\,] - E\ g(X)\ E\ h(X)\,] \geqslant 0$ (since X_1 and X_2 have the same distribution)

$\longleftrightarrow \mathrm{Cov}\ [\, g(X),\ h(X)\,] \geqslant 0.$

REFERENCES

Brumelle, S. and R. Vickson (1975) : "A unified approach to stochastic dominance" in "Stochastic optimization models in Finance", Ed. Ziemba-Vickson, Academic Press, 101-113

Chamberlain, G. (1979) : Heterogeneity, omitted variable bias and duration dependence, Harvard D.P.

Chesher, A. (1984) : "Testing for neglected heterogeneity", Econometrica 52, 865-872

Chesher, A. and T. Lancaster (1983) : "The estimation of models of labour market behaviour", R.E.S. 609-624

Congdon, P. (1985) : "Heterogeneity and timing effect in occupational mobility : a general model", Oxford Bulletin of Economics and Statistics 47, 347-369

Cox, D. and D. Oakes (1984) : "Analysis of survival data", Chapman-Hall

Elbers, C. and G. Ridder (1982) : "True and spurious duration dependence : the identifiability of the proportional hazard model", R.E.S. 49, 403-410

Feller, W. (1971) :"An introduction to probability theory and its applications", Vol. 2, New-York, Wiley, Chapter XIII

Flinn, C. and J.L Heckman (1982) : "New methods for analyzing structural models of labor force dynamics", Journal of Econometrics 18, 115-168

Heckman, J. (1981) : "Heterogeneity and state dependence", in Studies in labor markets, Ed. S. Rosen, University of Chicago Press

Heckman, J. and B. Singer (1984) : " Econometric duration analysis", Journal of Econometrics 24, 63-132

Kalbfleisch, J. and R. Prentice (1980) : "The statistical analysis of failure time data", Wiley

Kiefer, N. (1984) : "A simple test for heterogeneity in exponential models of duration", Journal of Labor Economics 2, 534-549

Lancaster, T. (1979) : "Econometric methods for the duration of unemployment" , Econometrica 47, 939-956

Lawless, J. (1982) : Statistical models and methods for lifetime data, Wiley

Rothschild, M. and J. Stiglitz (1970) : "Increasing risk I : A definition", JET 2, 225-247

Rothschild, M. and J. Stiglitz (1973) : "Some further results on the measurement of inequality", JET 6, 188-204

Spilerman, S. (1972) : "Extensions of the mover-stayer model", American Journal of Sociology, 78, 599-626.